ZooKeys 1135: 157–170 (2022)
doi: 10.3897/zookeys.1135.93631
https://zookeys.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

DNA barcoding of *Scomberomorus* (Scombridae, Actinopterygii) reveals cryptic diversity and misidentifications

Xiao-Shu Zeng[1*], Cheng-He Sun[1*], Xiao-Ying Huang[1], Ye-Ling Lao[1], Jin-Long Huang[1], Sha Li[2,3], Qun Zhang[1]

1 *Department of Ecology and Institute of Hydrobiology, Jinan University, Guangzhou 510632, China*
2 *Chinese Sturgeon Research Institute, China Three Gorges Corporation, Yichang 443100, Hubei, China*
3 *Hubei Key Laboratory of Three Gorges Project for Conservation of Fishes, Yichang 443100, Hubei, China*

Corresponding author: Qun Zhang (tqzhang@jnu.edu.cn)

Academic editor: Nina Bogutskaya | Received 18 August 2022 | Accepted 15 November 202 | Published 14 December 2022

https://zoobank.org/992FEA9A-5FF1-4DFA-B138-D91EC26EE6FD

Citation: Zeng X-S, Sun C-H, Huang X-Y, Lao Y-L, Huang J-L, Li S, Zhang Q (2022) DNA barcoding of *Scomberomorus* (Scombridae, Actinopterygii) reveals cryptic diversity and misidentifications. ZooKeys 1135: 157–170. https://doi.org/10.3897/zookeys.1135.93631

Abstract

The genus *Scomberomorus* is economically important; however, the taxonomic status and phylogenetic relationships in this genus are not clearly resolved, making it difficult to effectively protect and exploit fish resources. To clarify the taxonomic status of *Scomberomorus* species, mitochondrial cytochrome c oxidase I (COI) gene sequences of 150 samples were analyzed. The average genetic distance among 14 species was approximately 11 times greater than the distances within species, in accordance with the '10× rule' of species identification. Five of the 14 species did not form monophyletic clades based on a Bayesian inference gene tree. The application of four DNA-based species delimitation methods (automatic barcode gap discovery, barcode index numbers, Poisson tree process, and the K/θ method) yielded several key results. (1) Cryptic species were detected within *Scomberomorus commerson*. (2) A *Scomberomorus queenslandicus* sample from Australia was misidentified as *S. commerson* in the Barcode of Life Data System (BOLD). (3) Specimens originally identified as *Scomberomorus guttatus* was differentiated into four OTUs or species, two in the Yellow, South China, and Java seas, and two in geographically distant areas, one each in the Arabian Sea and the Bay of Bengal. (4) Six specimens from South Africa originally identified as *S. plurilineatus* most likely do not belong to the species. (5) Specimens identified as *S. maculatus* and *S. regalis* were conspecific; however, introgression cannot be ruled out. Our findings

* These authors contribute equally to this work.

Copyright *Xiao-Shu Zeng et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

revealed cryptic diversity and difficulties in morphological identification of species in the genus *Scomberomorus*. This study provides scientifically based support for the conservation of germplasm resources of the genus *Scomberomorus*.

Keywords
COI, conservation, cryptic species, DNA-based species delimitation, mackerel, phylogeny

Introduction

DNA barcoding provides a complementary approach to morphological species identification (Hebert et al. 2003a). The approximately 650 bp sequence at the 5' end of the animal cytochrome c oxidase I (COI) is a standard DNA barcoding region for delineating species (Hebert et al. 2003a; Vences et al. 2005). Hebert et al. (2003b, 2004) proposed the '10× rule' of species identification and concluded that intraspecific COI genetic distances are generally less than 2% based on an analysis of 13320 species in 11 phyla. Numerous studies (e.g., Zemlak et al. 2009; Pereira et al. 2013; Knebelsberger et al. 2015; Neves et al. 2020) have suggested that the COI gene is effective for differentiating among fish species and could be used for resolving synonymy, heteronomy, and identifying cryptic species.

The genus *Scomberomorus* belongs to the family Scombridae, one of the most popular and familiar food fish in the world (Yemmen and Gargouri 2022), composed of 18 species (Collette et al. 2001). They are rich in protein and highly unsaturated fatty acids and therefore possess high nutritional value (Lou et al. 2000). The biological characteristics and trends in the unit production of *Scomberomorus* fishes indicate that their resources are in a state of decline, and this can be attributed to overfishing and marine environmental pollution (Zheng et al. 2014). Therefore, research focused on the conservation of fish germplasm resources in the genus *Scomberomorus* is urgently needed.

Species identification in the genus *Scomberomorus* is mostly based on morphology (Collette and Russo 1985; Collette et al. 2001; Zhang et al. 2013a) and molecular data (e.g., Hoolihan et al. 2006; Habib and Sulaiman 2016; Mansourkiaei et al. 2016). Morphological identification mainly depends on the lateral line, pattern, and body color (Collette and Nauen 1983). However, the phenotype changes with growth, the color fades during preservation, and sexual dimorphism has been discovered (Collette and Nauen 1983) in some species of this genus, all of which make morphological identification difficult.

Previous molecular studies of the genus *Scomberomorus* have mostly focused on a few species within the genus (e.g., Habib and Sulaiman 2016; Mansourkiaei et al. 2016), and few studies (Banford et al. 1999; Jeena et al. 2022) have evaluated the whole genus. Owing to this lack of species representation, relationships within the *Scomberomorus* are unclear (Bayona-Vásquez et al. 2018). In this study, we conducted a DNA barcoding study of 150 specimens from the genus *Scomberomorus* to clarify their identification and provide scientific support for the conservation of germplasm resources.

Materials and methods

Ethical statement

The collection and sampling of specimens were reviewed and approved by the Animal Ethics Committee of Jinan University. All specimens used in this study were collected in accordance with Chinese laws. All experiments were performed to ensure optimal animal welfare and care.

Sample collection and morphological identification

Samples were collected from 11 locations in the coastal waters of China and 116 homologous sequences were downloaded from GenBank and the Barcode of Life Data System (BOLD). The collected information is shown in Fig. 1 and Suppl. material 1: table S1. All specimens preserved in 95% ethanol in Department of Ecology and Institute of Hydrobiology, Jinan University were identified based on morphological characters known in literature (Institute of Zoology et al. 1962; Collette and Russo 1985). For 116 homologous sequences downloaded from GenBank and BOLD, samples were morphologically identified according to the original publications (Suppl. material 1: table S1), excluding directly submitted sequences.

DNA extraction, amplification, and sequencing

Muscle tissue samples collected in our laboratory were extracted for the determination of COI sequences. DNA was extracted using a modified phenol/chloroform method (Le et al. 2010). PCR amplification was performed according to the method described by Ward et al. (2005). Universal primers FishF1 (5'-TCA ACC AAC CAC AAA GAC ATT GGC AC-3') and FishR1 (5'-TAG ACT TCT GGG TGG CCA AAG AAT CA-3') were used. The 20 μL polymerase chain reaction (PCR) mixtures contained 7 μL of sterilized ultrapure water, 10 μL of PCR Mix, 1 μL of each primer, and 1 μL of the DNA template. The reaction conditions were as follows: pre-denaturation at 95 °C for 5 min, denaturation at 94 °C for 40 s, annealing at 54 °C for 40 s, extension at 72 °C for 50 s, for 35 cycles, and a final extension at 72 °C for 10 min. PCR products were detected by 1% agarose gel electrophoresis, purified, and sequenced by BGI Genomics Co., Ltd. (Shenzhen, China). DNA extraction, amplification, and sequencing of the downloaded homologous sequences were also performed according to GenBank and BOLD.

Data analyses

The sequencing peaks were visualized using Chromas 2.6.6 (Technelysium Pty Ltd., South Brisbane, Queensland, Australia 2018) and the sequences were manually calibrated using Bioedit 7.2.5 (Hall 1999). Sequence characteristics were analyzed using

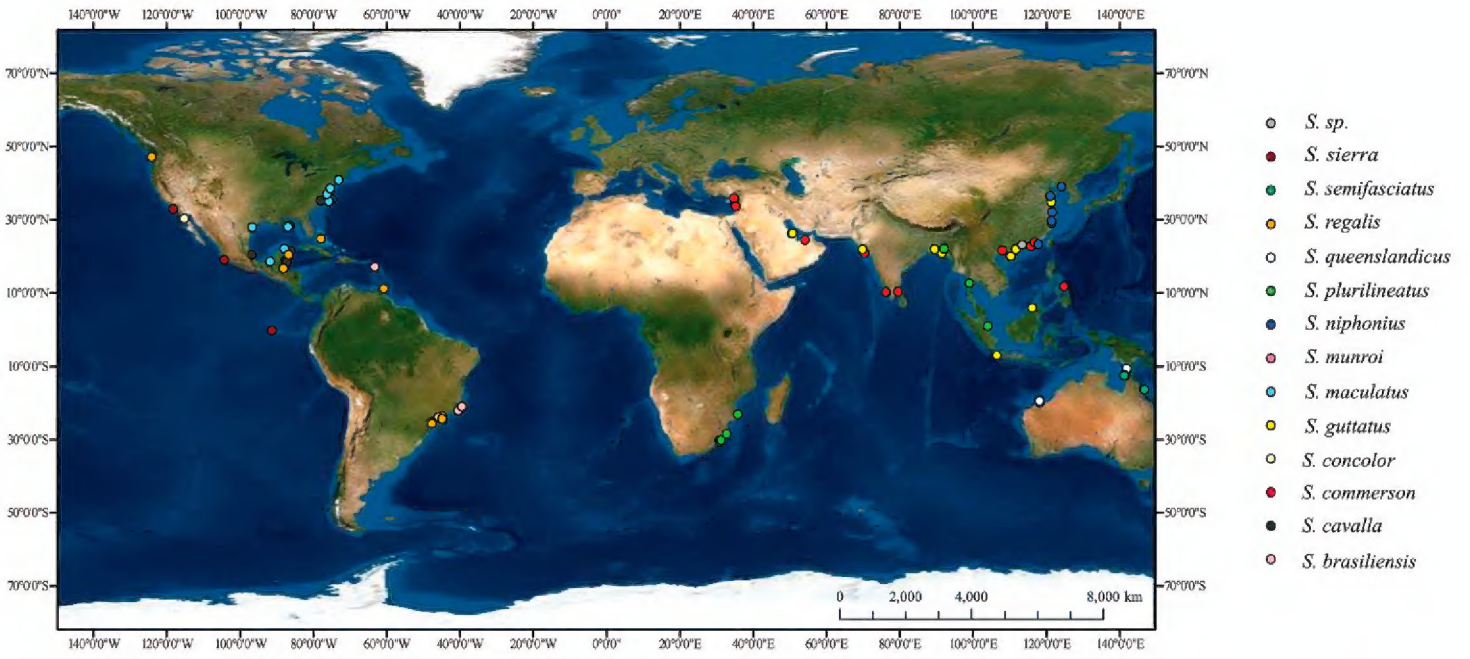

Figure 1. Localities of 150 samples in this study. One dot may represent more than one specimen.

MEGA 7.0 (Kumar et al. 2016), and various indices, such as base composition, variable sites (including parsimony-informative sites and singleton sites), and the transition-to-transversion ratio, were calculated.

All genetic distances were calculated based on Kimura two parameter (K2P) (Nei and Kumar 2000) distances using MEGA 7.0 (Kumar et al. 2016). A comparative analysis of all individuals within the same species was used to calculate the genetic distances between samples within each species, and these results were combined with the interspecific genetic distances to plot the barcode gap map of 14 species.

The COI gene sequences were tested for saturation using DAMBE 7.3.5 (Xia and Xie 2001). Based on the K2P substitution model, a neighbor-joining tree (NJ tree) was constructed, branch support was evaluated by 1000 repetitions of sampling, and genetic distances between and within clades were calculated. The construction of the Bayesian inference gene tree (BI tree) was performed using PhyloSuite v. 1.2.2 (Zhang et al. 2020). ModelFinder was used to select the best-fit partitioning model using the BIC criterion (Kalyaanamoorthy et al. 2017). BI phylogenies were inferred using MrBayes 3.2.6 (Ronquist et al. 2012) under the HKY+I+G+F model (two parallel runs, 2000000 generations), in which the initial 25% of sampled data were discarded as burn-in. FigTree v. 1.4.4 (Rambaut 2009) was used to visualize and edit the BI tree.

We employed four species delimitation methods: (1) Automatic Barcode Gap Discovery (ABGD) (Puillandre et al. 2012); (2) Barcode Index Numbers (BIN) (Ratnasingham and Hebert 2013) implemented in BOLD to obtain operational taxonomic units (OTUs); (3) Poisson tree processes (PTP) (Zhang et al. 2013b) implemented in the bPTP server (https://species.h-its.org/ptp/) with the BI tree as the input file; (4) the K/θ method: for species morphologically identified as conspecific, the mean pairwise distance within each clade (θ) and the minimum pairwise distance between clades (K) in the phylogenetic tree were recorded. Clades with K/θ ≥ 4 are considered reciprocally monophyletic with ≥ 95% probability (Birky et al. 2010).

Results

Sequence analysis

The COI gene sequences had an average length of 648 bp (567–652 bp) in 150 samples from 14 species in the genus *Scomberomorus*, and there were no base insertions or deletions. There were 453 conserved bases, accounting for 69.48% of the total number of bases, and 199 variable bases, accounting for 30.52% of the total number of bases, including 191 parsimony-informative sites and eight singleton bases. The A+T content (53.3%) was higher than the C+G content (46.7%), indicating an AT bias. The transition-to-transversion ratio was 3.1. A saturation analysis (Suppl. material 1: table S2) indicated that the base mutations did not reach saturation and were suitable for phylogenetic analyses.

Genetic distances and barcoding gaps

The intraspecific genetic distances of 14 species of the genus *Scomberomorus* were 0%–6.0%, with an average genetic distance of 1.18%. The interspecific genetic distances were 0.3%–17.4%, with an average genetic distance of 13.0%, which was approximately 11 times higher than estimates within species. Furthermore, 79% of intraspecific genetic distances were within the range of 0%–2%. The barcoding gap map (Fig. 2) showed that the maximum intraspecific genetic distances for *S. commerson*

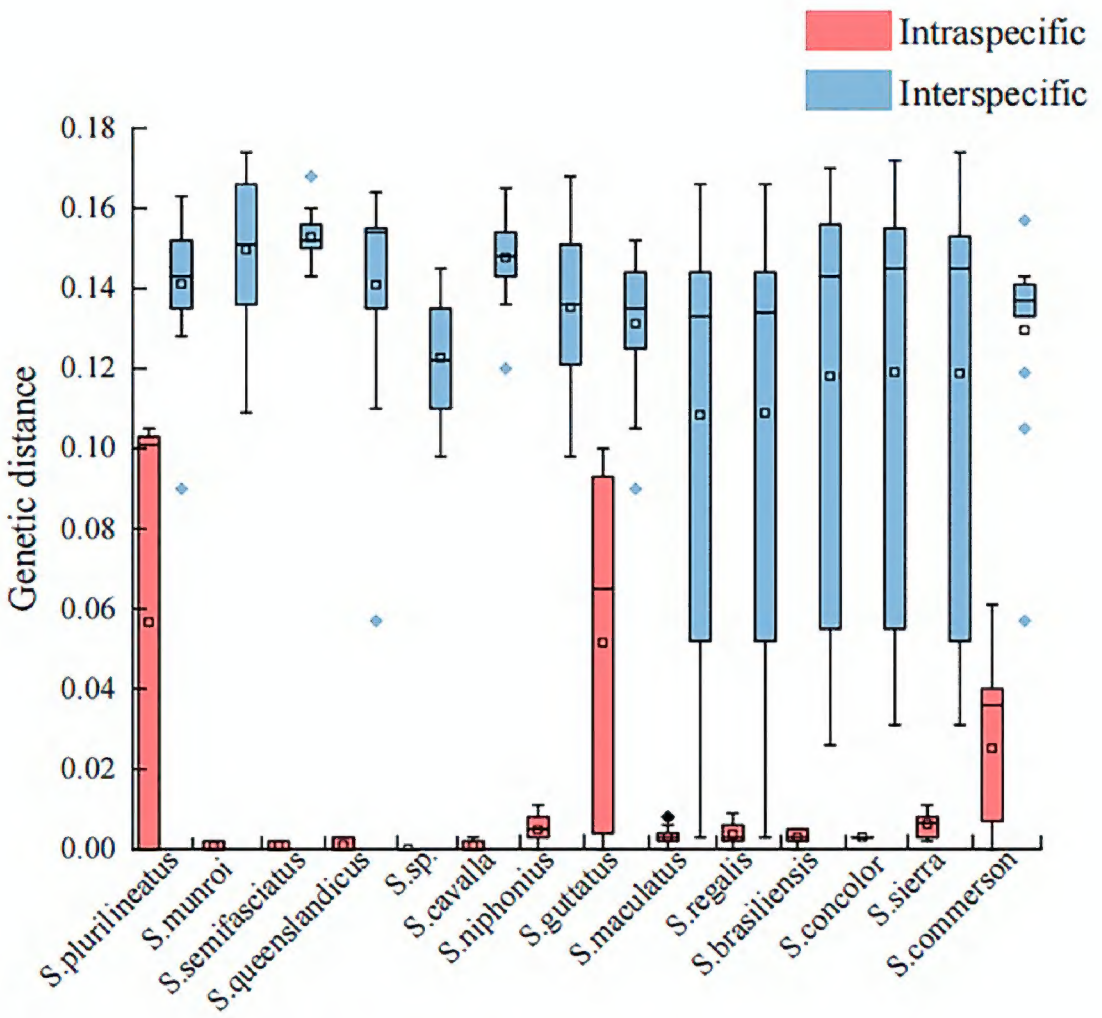

Figure 2. DNA barcoding gaps of 14 *Scomberomorus* species

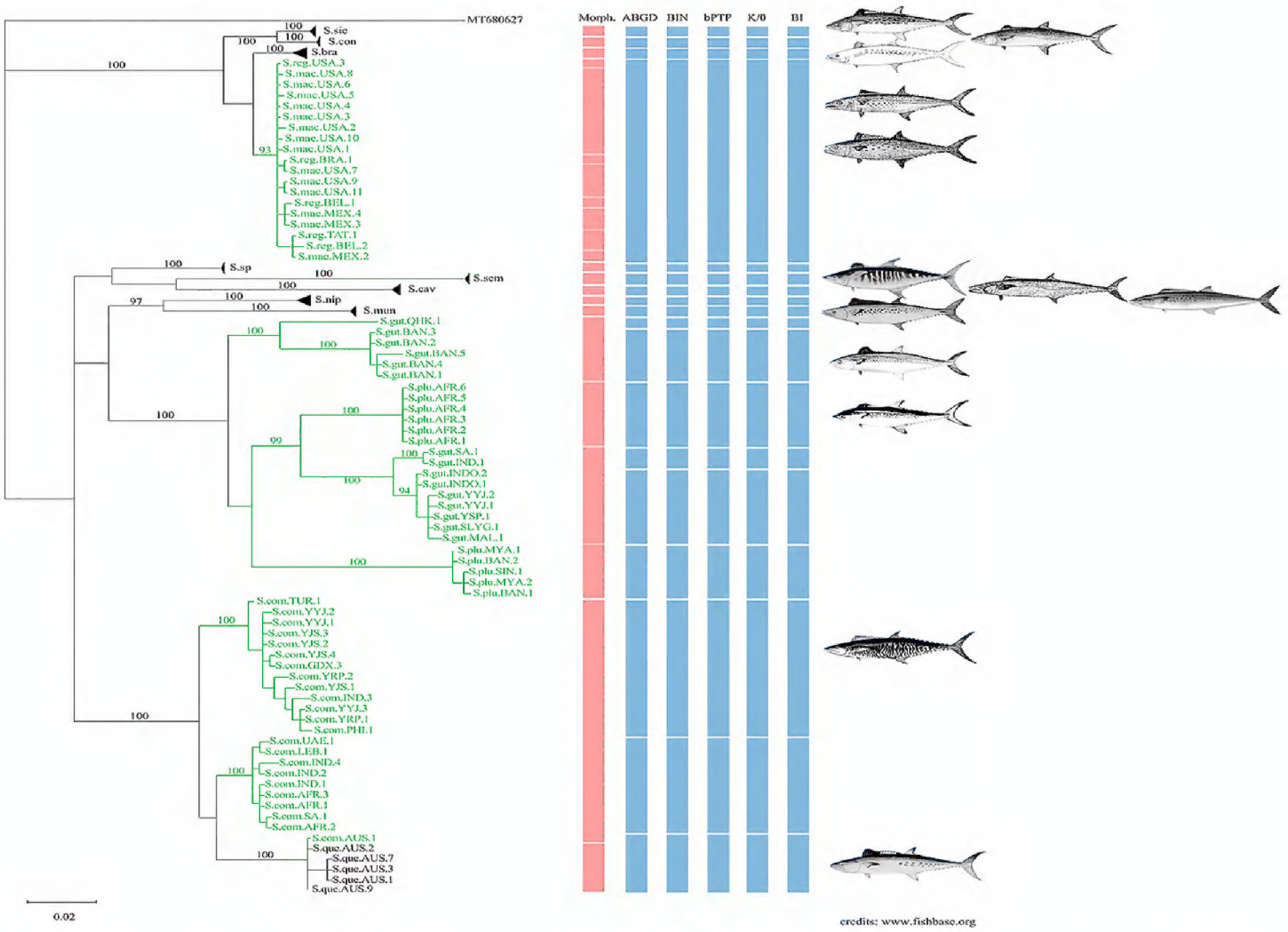

Figure 3. Bayesian inference (BI) tree based on the COI sequences of 14 *Scomberomorus* species. The green clades represent five species for which the species delimitation result is different based on morphology and the BI tree: *S. commerson*, *S. guttatus*, *S. plurilineatus*, *S. regalis*, and *S. maculatus*. Images of the genus *Scomberomorus* on the right from top to bottom are: *S. sierra*, *S. concolor*, *S. brasiliensis*, *S. regalis*, *S. maculatus*, *S. semifasciatus*, *S. cavalla*, *S. niphonius*, *S. munroi*, *S. guttatus*, *S. plurilineatus*, *S. commerson*, and *S. queenslandicus*. MT680627 is the outgroup. Numbers near the branches are bootstrap values.

(Lacepède, 1800), *S. guttatus* (Bloch & Schneider, 1801), *S. plurilineatus* Fourmanoir, 1966, *S. regalis* (Bloch, 1793), and *S. maculatus* (Mitchill, 1815) were not clearly different from the minimum interspecific genetic distances. In particular, the intraspecific genetic distances for *S. plurilineatus*, *S. guttatus*, and *S. commerson* were 6%, 5.2%, and 2.6%, respectively, and the interspecific genetic distance between *S. maculatus* and *S. regalis* was 0.3%. The remaining nine species had intraspecific distances of less than 2% and interspecific distances greater than 2%, forming clear DNA barcoding gaps.

Phylogenetic clustering analysis

A BI tree (Fig. 3) and a NJ tree (Suppl. material 1: fig. S1) exhibited similar topologies. All specimens in the BI tree formed 18 clades. The average genetic distance between clades was 12.97% (2.3%–17.4%), which was 48 times higher than the average genetic distance within clades of 0.27% (0%–0.75%). *S. commerson* formed two clades with 100 bootstrap values separated by a genetic distance of 3.9% in the BI tree. One *S. commerson* sample from Australia was assigned to a lineage with *S. plurilineatus*

samples from Australia. *S. plurilineatus* and *S. guttatus* clustered together in one large clade, and *S. guttatus* formed three clades with 100, 100, and 94 bootstrap values in the BI tree. Together with the sample from Haikou, Hainan, China, *S. guttatus* was divided into four small clades with an inter-clade genetic distance of 7.48%, which was 31 times higher than the average intra-clade genetic distance (0.24%). *S. maculatus* formed a clade with *S. regalis* with 93 bootstrap value in the BI tree, and the genetic distance within the clade was only 0.3%. All other species formed monophyletic groups.

DNA-based species delimitation

Four DNA-based species delimitation methods yielded consistent results (Fig. 3). All methods supported the division of 14 species into 18 hypothetical species by classifying *S. commerson* into two hypothetical species, *S. guttatus* into four hypothetical species, and *S. plurilineatus* into two hypothetical species and combining *S. regalis* with *S. maculatus*.

In the ABGD analysis (Suppl. material 1: fig. S2), a good barcode gap was observed when a priori intraspecific divergence was 0.0046, and this barcode gap strongly supported the division of taxa into 18 groups. The BIN analysis (Suppl. material 1: table S3) divided the 150 samples into 18 OTUs in which the genetic distance between the OTU and the nearest OTU (NN Dist) was greater than the internal maximum genetic distance (Max), indicating apparent divergence. In the bPTP analysis (Table 1), combined with the results based on morphological characters, the divisions based on bPTP (ML) were selected as the species definitions. A map (Fig. 4) of new clades of *S. guttatus* (OTU-10, OTU-11, OTU-13, and OTU-14) and *S. commerson* (OTU-16 and OTU-17) was obtained based on the bPTP results. In the K/θ method (Suppl. material 1: table S4), the clades with $\theta = 0$ and the clade with only one sample (clade 10 in the BI tree) were considered separate OTUs; therefore, 18 OTUs were obtained.

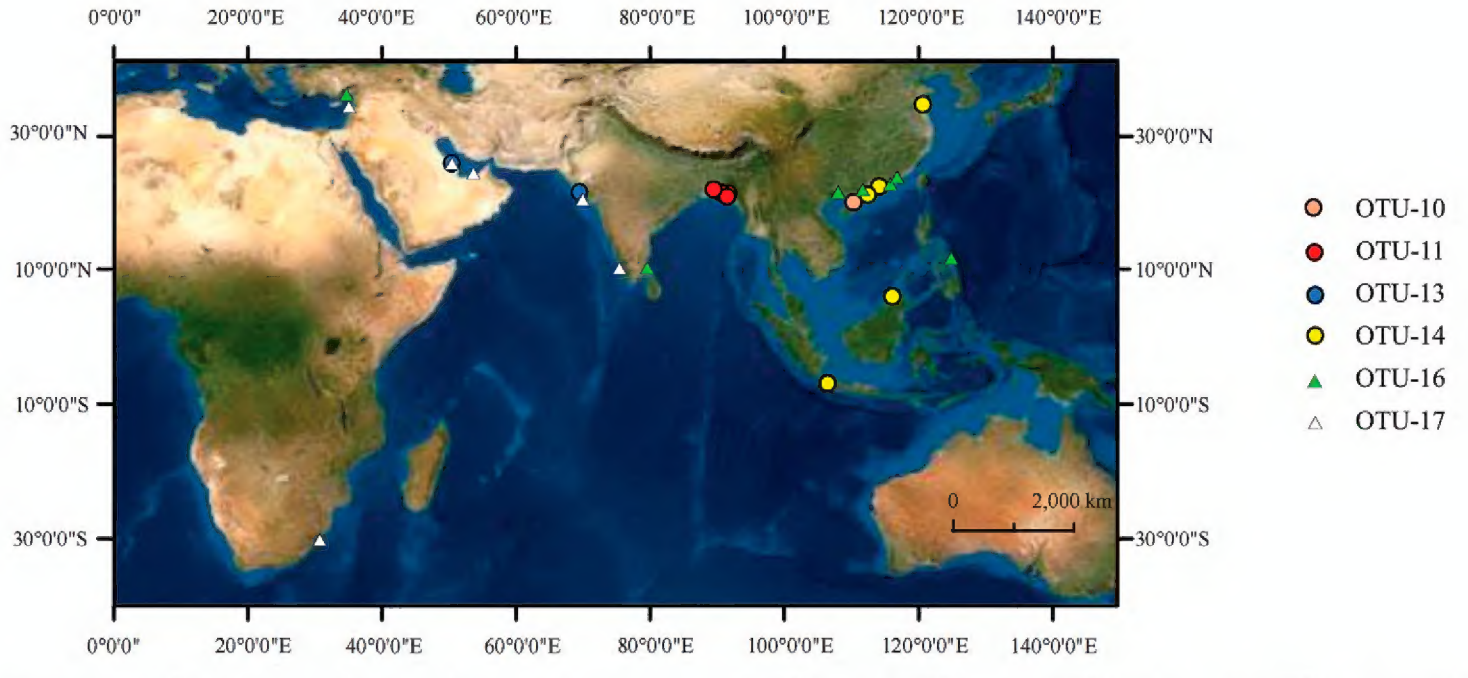

Figure 4. Localities of new clades of *S. guttatus* (OTU-10, OTU-11, OTU-13, and OTU-14) and *S. commerson* (OTU-16 and OTU-17) based on the bPTP results.

Table 1. Results for 14 *Scomberomorus* species based on a bPTP analysis.

Morphological species	OTU	Number	Catalog number
S. sierra	OTU-1	4	S.sie.ECU.1, S.sie.MEX.1–2, S.sie.USA.1
S. concolor	OTU-2	2	S.con.MEX.1, S.con.MEX.2
S. brasiliensis	OTU-3	14	S.bra.NET.1, S.bra.BRA.1–13
S. maculatus and *S. regalis*	OTU-4	28	S.mac.USA.1–11, S.mac.MEX.1–5, S.reg.BRA.1–3, S.reg.BEL.1–3, S.reg.TAT.1, S.reg.BAH.1, S.reg.MEX.1, S.reg.USA.1–3
S. sp	OTU-5	3	S.sp.YSP.1–3
S. semifasciatus	OTU-6	5	S.sem.AUS.1–5
S. cavalla	OTU-7	12	S.cav.USA.1–2, S.cav.MEX.1–7, S.cav.ALT.1, S.cav.USA.3–4
S. niphonius	OTU-8	14	S.nip.SLS.1–4, S.nip.LYK.1–3, S.nip.YRP.1–2, S.nip.LDG.1–3, S.nip.ZSM.1, S.nip.ZZS.1
S. munroi	OTU-9	6	S.mun.AUS.1–6
S. guttatus	OTU-10	1	S.gut.QHK.1
	OTU-11	5	S.gut.BAN.1–5
	OTU-12	6	S.plu.AFR.1–6
	OTU-13	2	S.gut.IND.1, S.gut.SA.1
S. plurilineatus	OTU-14	9	S.gut.INDO.1–2, S.gut.MAL.1, S.gut.SCS.1, S.gut.SLYG.1, S.gut.YSP.1–2, S.gut.YYJ.1–2
	OTU-15	5	S.plu.MYA.1–2, S.plu.BAN.1–2, S.plu.SIN.1
S. commerson	OTU-16	15	S.com.IND.3, S.com.PHI.1, S.com.GDX.1–3, S.com.YYJ.1–3, S.com.YRP.1–2, S.com.YJS.1–4, S.com.TUR.1
	OTU-17	9	S.com.AFR.1–3, S.com.SA.1, S.com.IND.1–2, S.com.IND.4, S.com.LEB.1, S.com.UAE.1
S. queenslandicus	OTU-18	10	S.com.AUS.1, S.que.AUS.1–9

Discussion

According to the '10× rule' of species identification and 2% threshold (Hebert et al. 2003b, 2004), taxonomic uncertainty was discovered in the samples of *S. commerson*, *S. guttatus*, *S. plurilineatus*, *S. maculatus*, and *S. regalis* on the basis of genetic distances. Phylogenetic trees and four DNA molecular definition methods all support the classification of 14 morphological species into 18 hypothetical species.

Taxonomic identification of *S. commerson* and *S. queenslandicus* individuals

Intraspecific genetic distance in *S. commerson* (2.6%) was slightly greater than the threshold of 2%; *S. commerson* samples were assigned to three lineages in the phylogenetic trees. The samples from Australia were mixed with *S. queenslandicus* (Munro, 1943) on a single clade. Both species are distributed in Australia (Collette and Nauen 1983; Collette et al. 2001; Kuiter 2021). The adult color pattern in the two species differ: *S. commerson* has many thin, wavy, vertical stripes on the side of the body, while *S. queenslandicus* adults have about three indistinct rows of bronze-grey blotches on the sides (Collette and Nauen 1983). Adult *S. commerson* (generally 120 cm in length) is larger than *S. queenslandicus* (generally 80 cm in length), and juvenile *S. commerson* often have blotches. Therefore, we assumed that *S. commerson* from Australia in the BOLD was misidentified and was actually *S. queenslandicus*; however, further identi-

fication and determination were impossible because the corresponding image was not available in the database.

The bootstrap values for the other two clades of *S. commerson* were 100 in the BI tree. The samples in this study and the three samples from India, the Philippines, and Turkey were assigned to the same clade, and the remaining samples belonged to a separate clade. The two clades did not show obvious geographical clustering (Fig. 4), and the genetic distance between clades was 3.9% (greater than the 2% threshold). Vineesh et al. (2018) detected at least three genetically distinct populations of *S. commerson* in the Indian Ocean region. Johnson et al. (2021) studied the genetic population structure and phylogenetic relationships in the coastal waters of northern Tanzania and concluded that *S. commerson* is a single mixed population with high genetic diversity. In this study, the splitting of *S. commerson* into two clades in the molecular phylogenetic trees without obvious geographical clustering may be explained by the migratory behavior of the mackerels and the easy dispersal of their larvae over a wide area by ocean currents (Hoolihan et al. 2006), leading to secondary contact after differentiation. Therefore, cryptic species may exist in this lineage.

Taxonomic identification of *S. guttatus* and *S. plurilineatus* individuals

The intraspecific genetic distance in *S. guttatus* was 5.2%, which was significantly greater than the 2% threshold. Four clades were formed in the phylogenetic trees, and the genetic distances between the clades were greater than 2% in the phylogenetic trees. The inter-clade genetic distance (7.48%) was 31 times higher than the average intra-clade genetic distance (0.24%), which was in accordance with the '10× rule' of species identification (Hebert et al. 2003b, 2004). Except for the sample from Haikou, Hainan, China, the other three clades clustered according to geographical distributions. The first clade consisted of individuals from the Bay of Bengal, the second clade consisted of individuals from the Arabian Gulf and Arabian Sea, and the third clade consisted of individuals from the Yellow Sea, South China Sea, and Java Sea (Fig. 4). Yu et al. (2021) conducted a DNA barcoding analysis of *Jaydia smithi* (Kotthaus, 1970) and found that the Chinese and Mediterranean populations could be divided into two groups, with an average genetic distance between the two groups of 0.044, suggesting the presence of cryptic species. Chen et al. (2017) found that the genetic distance between two *Terapon puta* (Cuvier, 1829) groups was 5% in the western Pacific and Indo-Mediterranean and suggested that the species might be divided into two subspecies or even two species. Luo et al. (2021) found that the average intraspecific genetic distance of *Lateolabrax* spp. was 3.91%, and two clades corresponding to populations in China and Japan were found in the NJ tree; the average inter-clade genetic distance (6.98%) was 14.2 times higher than the average intra-clade genetic distance (0.49%), supporting the division into two species, *L. japonicus* and *L. maculatus*. A study of *S. guttatus* from the South China Sea (Ye 2012) revealed that 19 individuals formed two major clades and suggested that the group originated from two different maternal ancestors. These findings are highly similar to the results of the present study, in which samples from the South China Sea clustered into two clades. Our data indicate that *S. guttatus* complex involves four OTUs or even four species.

S. plurilineatus formed two distant clades in the phylogenetic trees, one of which mixed with *S. guttatus*. the name *S. guttatus* has been misapplied to *S. plurilineatus* (Collette and Nauen 1983); therefore, six specimens of *S. plurilineatus* from South Africa might have been misidentified. According to a previous publication (Steinke et al. 2016), three specimens (JF494458–JF494460) were adults and morphological identification was performed; however, specimen images were not available in GenBank or BOLD. The other three specimens (DSLAG600-10, DSLAG1283-11, and DSLAG1287-11) were larvae and images of juvenile and fish eggs were obtained; however, it was not possible to accurately identify them based on the images alone.

Taxonomic identification of *S. maculatus* and *S. regalis* individuals

The interspecific genetic distance between *S. maculatus* and *S. regalis* was 0.3% and the taxa were mixed on a single clade of the phylogenetic trees. According to FishBase (https://fishbase.org/) and the localities of samples in this study (Fig. 1), *S. maculatus* and *S. regalis* are both distributed in the western Atlantic Ocean, and *S. regalis* has a wider range. Banford et al. (1999) found little difference between the mitochondrial genomes of *S. maculatus* and *S. regalis* and hypothesized that the hybridization with *S. regalis* resulted in the introgressive loss of the *S. maculatus* mtDNA genome. Based on the 2% threshold and with reference to the geographic distributions of the two species, we speculated that *S. maculatus* and *S. regalis* are the same species. However, it is possible that introgressive hybridization affected the results, and we cannot exclude the possibility that the two are actually separate species.

The discovery of cryptic species in this study expands current estimates of biodiversity and allows better precautionary and scientifically management, which is important to plan reasonable conservation strategies (Loxdale et al. 2016). However, owing to the wide geographical distribution of the samples, it was difficult to obtain representative samples for comparison, leading to the potential for species misidentification. In addition, the maternally inherited mitochondrial COI does not reflect bi-parentally inherited nuclear genome information, making it difficult to distinguish introgressive hybridization. Therefore, future studies of species identification should combine morphometric, nuclear genetic markers, and biological analyses to further clarify the taxonomic status of the genus *Scomberomorus* so as not to destroy the available resources (Wuketits 1997).

Acknowledgements

We would like to thank Editage (www.editage.cn) for English language editing.

The present study was supported by the National Key R&D Program of China (Grant number 2018YFD0900802); Fishery resources survey of Guangxi Zhuang Autonomous Region (GXZC2022-G3-001062-ZHZB); Director's Fund of the Hubei Key Laboratory of Three Gorges Project for Conservation of Fishes, China Three Gorges Corporation (0704157) and the Outstanding Innovative Talents Cultivation Funded Programs for Doctoral Students of Jinan University (Project number 2021CXB022).

References

Banford HM, Bermingham E, Collette BB, McCafferty SS (1999) Phylogenetic systematics of the *Scomberomorus regalis* (Teleostei: Scombridae) species group: molecules, morphology and biogeography of Spanish mackerels. Copeia 1999(3): 596–613. https://doi.org/10.2307/1447593

Bayona-Vásquez NJ, Glenn TC, Domínguez-Domínguez O, Uribe-Alcocer M, Díaz-Jaimes P (2018) Mitochondrial genomes of the Pacific sierra mackerel *Scomberomorus sierra* and the Monterey Spanish mackerel *Scomberomorus concolor* (Perciformes, Scombridae). Conservation Genetics Resources 10(3): 471–474. https://doi.org/10.1007/s12686-017-0851-9

Birky Jr CW, Adams J, Gemmel M, Perry J (2010) Using population genetic theory and DNA sequences for species detection and identification in asexual organisms. PLoS ONE 5(5): e10609. https://doi.org/10.1371/journal.pone.0010609

Chen L, Huang BY, Xue D (2017) DNA barcoding of Terapontidae and taxonomic status of a new record of *Terapon puta* in coastal waters of China. Marine Fisheries 39(2): 121–130. https://doi.org/10.13233/j.cnki.mar.fish.2017.02.001 [in Chinese]

Collette BB, Nauen CE (1983) An annotated and illustrated catalogue of Tunas, Mackerels, Bonitos and related species known to date. FAO Species Catalogue. Vol. 2. Scombrids of the world. FAO Fisheries Synopsis 125(2).

Collette BB, Russo JL (1985) Morphology, systematics, and biology of the Spanish Mackerels (*Scomberomorus*, Scombridae). Fish Bulletin 82(4): 545–692.

Collette BB, Reeb C, Block BA (2001) Systematics of the tunas and mackerels (Scombridae). Fish Physiology 19: 1–33. https://doi.org/10.1016/S1546-5098(01)19002-3

Habib A, Sulaiman Z (2016) High genetic connectivity of narrow-barred Spanish mackerel (*Scomberomorus commerson*) from the South China, Bali and Java Seas. Zoology and Ecology 26(2): 93–99. https://doi.org/10.1080/21658005.2016.1161121

Hall TA (1999) Bioedit: A user-friendly biological sequence alignment program for windows 95/98/nt. Nucleic Acids Symposium Series 41(41): 95–98. https://doi.org/10.1021/bk-1999-0734.ch008

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003a) Biological identifications through DNA barcodes. Proceedings of the Royal Society of London, Series B, Biological Sciences 270(1512): 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert PDN, Ratnasingham S, DeWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London, Series B: Biological Sciences 270(suppl_1): S96–S99. https://doi.org/10.1098/rsbl.2003.0025

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM, Godfray C (2004) Identification of birds through DNA barcodes. PLoS Biology 2(10): e312. https://doi.org/10.1371/journal.pbio.0020312

Hoolihan JP, Anandh P, van Herwerden L (2006) Mitochondrial DNA analyses of narrow-barred Spanish mackerel (*Scomberomorus commerson*) suggest a single genetic stock in the ROPME sea area (Arabian Gulf, Gulf of Oman, and Arabian Sea). ICES Journal of Marine Science 63(6): 1066–1074. https://doi.org/10.1016/j.icesjms.2006.03.012

Institute of Zoology, Chinese Academy of Sciences, Institute of Oceanology, Chinese Academy of Sciences, Shanghai Ocean University (1962) Fish of the South China Sea. Science Press, Beijing. [in Chinese]

Jeena NS, Rahuman S, Roul SK, Azeez PA, Vinothkumar R, Manas HM, Nesnas EA, Rathinam AMM, Surya S, Rohit P, Abdussamad EM, Gopalakrishnan A (2022) Resolved and Redeemed: A New Fleck to the Evolutionary Divergence in the Genus *Scomberomorus* Lacépède, 1801 (Scombridae) With Cryptic Speciation. Frontiers in Marine Science 9: e888463. https://doi.org/10.3389/fmars.2022.888463

Johnson MG, Mgaya YD, Shaghude YW (2021) Analysis of the genetic stock structure and phylogenetic relationship of narrow-barred Spanish mackerel *Scomberomorus commerson* (Lacépède, 1800) along the northern Tanzanian coastal waters using mitochondrial DNA. Regional Studies in Marine Science 46: e101862. https://doi.org/10.1016/j.rsma.2021.101862

Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. Nature Methods 14(6): 587–589. https://doi.org/10.1038/nmeth.4285

Knebelsberger T, Dunz AR, Neumann D, Geiger MF (2015) Molecular diversity of Germany's freshwater fishes and lampreys assessed by DNA barcoding. Molecular Ecology Resources 15(3): 562–572. https://doi.org/10.1111/1755-0998.12322

Kuiter RH (2021) A Guide to Sea Fishes of Australia. Reed New Holland, Chatswood.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution 33(7): 1870–1874. https://doi.org/10.1093/molbev/msw054

Le XL, Zhang Q, Zhao S (2010) A fast and efficient method for isolation of genomic DNA from fish specimens. Shengwu Jishu Tongbao 20(2): 202–204. https://doi.org/10.13560/j.cnki.biotech.bull.1985.2010.02.014 [in Chinese]

Lou YJ, Yang WG, Wang Y, Qiu DH, Dong MM (2000) Development of Spanish mackerel convenience food. Food Science and Technology 25(06): 31. https://doi.org/10.13684/j.cnki.spkj.2000.06.017 [in Chinese]

Loxdale HD, Davis BJ, Davis RA (2016) Known knowns and unknowns in biology. Biological Journal of the Linnean Society 117(2): 386–398. https://doi.org/10.1111/bij.12646

Luo C, Zhang Q, Huang ZJ, Zhang YN (2021) DNA barcoding of some common coastal fishes from Japan and China. Marine Fisheries 43(01): 1–11. https://doi.org/10.3724/SP.J.1004-2490.2021.0101 [in Chinese]

Mansourkiaei A, Mostafavi PG, Fatemi SMR, Kaymaram F, Nazemi A (2016) Phylogenetic relationships of *Scomberomorus commerson* using sequence analysis of the mtDNA D-loop region in the Persian Gulf, Oman Sea and Arabian Sea. international aquatic research 8(2): 137–148. https://doi.org/10.1007/s40071-016-0129-y

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, USA.

Neves JM, Almeida JP, Sturaro MJ, Fabre NN, Pereira RJ, Mott T (2020) Deep genetic divergence and paraphyly in cryptic species of *Mugil* fishes (Actinopterygii: Mugilidae). Systematics and Biodiversity 18(2): 116–128. https://doi.org/10.1080/14772000.2020.1729892

Pereira LH, Hanner R, Foresti F, Oliveira C (2013) Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? BMC Genetics 14(1): 1–14. https://doi.org/10.1186/1471-2156-14-20

Puillandre N, Lambert A, Brouillet S, Achaz G (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular Ecology 21(8): 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Rambaut A (2009) FigTree version 1.4.4. Computer program distributed by the author. http://tree.bio.ed.ac.uk/software/figtree/

Ratnasingham S, Hebert PDN (2013) A DNA-based registry for all animal species: The Barcode Index Number (BIN) system. PLoS ONE 8(7): e66213. https://doi.org/10.1371/journal.pone.0066213

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61(3): 539–542. https://doi.org/10.1093/sysbio/sys029

Steinke D, Connell AD, Hebert PD (2016) Linking adults and immatures of South African marine fishes. Genome 59(11): 959–967. https://doi.org/10.1139/gen-2015-0212

Technelysium Pty Ltd., South Brisbane, Queensland, Australia (2018) Chromas version 2.6.6.

Vences M, Thomas M, Bonett RM, Vieites DR (2005) Deciphering amphibian diversity through DNA barcoding: Chances and challenges. Philosophical Transactions of the Royal Society B: Biological Sciences 360(1462): 1859–1868. https://doi.org/10.1098/rstb.2005.1717

Vineesh N, Divya PR, Kathirvelpandian A, Mohitha C, Shanis CPR, Basheer VS, Gopalakrishnan A (2018) Four evolutionarily significant units among narrow-barred Spanish mackerel (*Scomberomorus commerson*) in the Indo-West Pacific region. Marine Biodiversity 48(4): 2025–2032. https://doi.org/10.1007/s12526-017-0714-3

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B: Biological Sciences 360(1462): 1847–1857. https://doi.org/10.1098/rstb.2005.1716

Wuketits FM (1997) The status of biology and the meaning of biodiversity. Naturwissenschaften 84(11): 473–479. https://doi.org/10.1007/s001140050431

Xia X, Xie Z (2001) DAMBE: Software package for data analysis in molecular biology and evolution. The Journal of Heredity 92(4): 371–373. https://doi.org/10.1093/jhered/92.4.371

Ye HW (2012) The South China Sea Spotted Mackerel Population Genetic Diversity Research. Beijing Agriculture (12): 138–139. [in Chinese]

Yemmen C, Gargouri M (2022) Potential hazards associated with the consumption of Scombridae fish: Infection and toxicity from raw material and processing. Journal of Applied Microbiology 132(6): 4077–4096. https://doi.org/10.1111/jam.15499

Yu ZS, Song N, Motomura H, Gao T (2021) Taxonomic revision of the cardinalfish genus *Jaydia* in China. Biodiversity Science 29(7): 971–979. https://doi.org/10.17520/biods.2020320 [in Chinese]

Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PDN (2009) DNA barcoding reveals overlooked marine fishes. Molecular Ecology Resources 9: 237–242. https://doi.org/10.1111/j.1755-0998.2009.02649.x

Zhang C, Ye Z, Panhwar SK, Shen W (2013a) Stock discrimination of the Japanese Spanish mackerel (*Scomberomorus niphonius*) based on the otolith shape analysis in the Yellow Sea and Bohai Sea. Journal of Applied Ichthyology 29(2): 368–373. https://doi.org/10.1111/jai.12084

Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013b) A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29(22): 2869–2876. https://doi.org/10.1093/bioinformatics/btt499

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT (2020) PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular Ecology Resources 20(1): 348–355. https://doi.org/10.1111/1755-0998.13096

Zheng YJ, Li JS, Zhang QY, Hong WS (2014) Research progresses of resource biology of important marine pelagic food fishes in China. Shuichan Xuebao (01): 149–160. https://doi.org/10.3724/SP.J.1231.2014.48799 [in Chinese]

Supplementary material 1

DNA barcoding of *Scomberomorus* (Scombridae, Actinopterygii) reveals cryptic diversity and misidentifications

Authors: Xiao-Shu Zeng, Cheng-He Sun1, Xiao-Ying Huang, Ye-Ling Lao, Jin-Long Huang, Sha Li, Qun Zhang

Data type: Data and image (word file).

Explanation note: Specimen information of *Scomberomorus* in this study. Test of the substitution saturation of COI gene sequences of the genus *Scomberomorus*. Results of 14 *Scomberomorus* species based on the BIN analysis. Results of 14 *Scomberomorus* species based on the K/θ method. NJ tree based on the COI sequences of 14 *Scomberomorus* species. The green clades represent five species for which the species delimitation result is different based on morphology and the NJ tree: *S. commerson*, *S. guttatus*, *S. plurilineatus*, *S. regalis*, and *S. maculatus*. MT680627 is the outgroup. Numbers near the branches are bootstrap values. ABGD analysis of the genus *Scomberomorus*.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.1135.93631.suppl1